

CLAIMS

1. A system for providing a service to identify and structure free text databases, comprising:
 - a processing element;
 - one or more data access ports, said ports providing access to data by said processing element;
 - a set of one or more input devices readable by said processing element;
 - a storage device, said storage device containing instructions executable by said processing element to perform the functions of:
 - (i) reading an access reference through said input device set, the access reference referencing a customer's source of unstructured data, the unstructured data including free text;
 - (ii) accessing the source of unstructured data;
 - (iii) identifying text records within the free text;
 - (iv) linguistically parsing the identified text records;
 - (v) identifying roles within the parsed text records, said identifying producing attribute extractions, each of said attribute extractions containing attribute information of the derived source text;
 - (v) applying caseframes to the attribute extractions, said applying caseframes producing a filtered set of attributes extractions; and
 - (vi) producing a structured data element containing the filtered set of extractions.
2. A system according to claim 1, wherein said instructions are further executable to perform the function of identifying domains of the filtered set of attributes extractions.
3. A system according to claim 2, wherein said instructions are further executable to perform the function of the assignment of domain roles.
4. A system according to claim 3, wherein said instructions are further executable to perform the function of producing relation types.
5. A system according to claim 4, wherein said instructions are further executable to perform the function of coalescing the produced relation types.

6. A system according to claim 1, wherein said instructions are further executable to perform the function of creating a new database containing the structured data element produced in said producing.
7. A system according to claim 1, wherein the instructions are further executable to produce a new relational database containing the structured data element produced in said producing a structured data element.
8. A system according to claim 1, wherein said instructions are further executable to perform the function of producing a file containing the structured data element produced in said producing a structured data element.
9. A system according to claim 8, wherein the instructions are further executable to produce a file having a format containing the structured data element selected from the group of XML, character separated values, spreadsheet formats and file-based database structures.
10. A system according to claim 1, wherein the structured data element produced by the performance of said producing includes reference information to the original free text for construed data.
11. A system according to claim 1, wherein the instructions are further executable to construct a library containing extracted attributes.
12. A system according to claim 1, wherein the instructions are further executable to construct a library containing extracted relational facts.
13. A system according to claim 1, wherein said instructions are further executable to perform the function of combining like relation types.
14. A system according to claim 1, wherein said instructions are further executable to perform the

function of combining like attributes.

15. A system according to claim 1, wherein said instructions are further executable to perform the function of coalescing the identified attributes.

16. A system according to claim 1, wherein said instructions are further executable to perform the function of coalescing identified relation types.

17. A method for providing a service to identify and structure free text databases, comprising the steps of:

- reading an access reference through said input device set, the access reference referencing a customer's source of unstructured data, the unstructured data including free text;
- accessing the source of unstructured data;
- identifying text records within the free text;
- linguistically parsing the identified text records;
- identifying roles within the parsed text records, said identifying producing attribute extractions, each of said attribute extractions containing attribute information of the derived source text;
- applying caseframes to the role extractions, said applying caseframes producing a filtered set of attribute extractions; and
- producing a structured data element containing the filtered set of extractions.

18. A method according to claim 17, further comprising the step of identifying domains of the filtered set of attribute extractions.

19. A system according to claim 18, further comprising the step of assigning domain roles.

20. A system according to claim 19, further comprising the step of producing relation types.

21. A system according to claim 20, further comprising the step of coalescing the produced relation types.

22. A system according to claim 18, further comprising the step of creating a new database containing the structured data element produced in said producing.
23. A system according to claim 18, further comprising the step of producing a new relational database containing the structured data element produced in said producing a structured data element.
24. A system according to claim 18, further comprising the step of producing a file containing the structured data element produced in said producing a structured data element.
25. A system according to claim 24, further comprising the step of producing a file having a format containing the structured data element selected from the group of XML, character separated values, spreadsheet formats and file-based database structures.
26. A system according to claim 18, wherein the structured data element produced by the performance of said producing includes reference information to the original free text for construed data.
27. A system according to claim 18, further comprising the step of constructing a library containing extracted attributes.
28. A system according to claim 18, further comprising the step of constructing a library containing extracted relational facts.
29. A system according to claim 19, further comprising the step of combining like relation types.
30. A system according to claim 18, further comprising the step of combining like attributes.
31. A system according to claim 18, further comprising the step of coalescing the identified attributes.

32. A system according to claim 18, further comprising the step of coalescing identified relation types.